# FAST MATRIX MULTIPLICATION*

Charles M. Fiduccia
Department of Computer Science
State University of New York
at Stony Brook
Stony Brook, New York 11790 U.S.A.



Proceedings of the Third Annual ACM Symposium on Theory of Computing, May 1971.

*This research was supported by the National Aeronautics and Space Administration grant numbers NGR40-002-082 and NGR40-002-090.

Not in 70, 71, 72

# FAST MATRIX MULTIPLICATION

C. M. Fiduccia

Center for Computer and Information Sciences
Brown University
Providence, Rhode Island

Presented at the Third ACM Symposium on the Theory of Computing.

This research was supported by the National Aeronautics and Space Administration, contract numbers NGR 40-002-082 and NGR 40-002-090.

## INTRODUCTION

Given the entries of an m×n matrix A and those of a column n-vector b, the entries of the product Ab can be computed using mn multiplications and m(n-1) additions by direct application of the formula

$$(Ab)_i = a_{i1}b_1 + \ldots + a_{in}b_n \qquad i = 1,\ldots,m.$$

However, in many cases the matrix A has a particular form, and Ab can be computed with fewer operations. For example, the finite Fourier transform $y_0,\ldots y_{n-1}$ of $x_0,\ldots,x_{n-1}$

$$y_r = \sum_{s=0}^{n-1} \omega^{rs} x_s \qquad \omega = e^{2\pi i/n}$$

can be computed with about $n\log_2 n$ multiplications and $n\log_2 n$ additions by using the fast Fourier transform (FFT) algorithm [4].

This paper deals with three aspects of algebraic complexity. The first section is concerned with lower bounds on the number of operations required to compute several functions. Several theorems are presented and their proofs sketched. The second section deals with relationships among the complexities of several sets of functions. In the third section, several matrices of general interest are examined and upper bounds on the number of operations required to multiply by them are constructively derived.

## LOWER BOUNDS

It is sufficient to consider the product Ab of a matrix A and a vector b, since computing the entries of the product BC of an m×n matrix B and an n×p matrix C is equivalent to computing the entries of Ab, where

$$A = \begin{pmatrix} B & & & \\ & B & & \\ & & \ddots & \\ & & & B \end{pmatrix} \qquad b = \begin{pmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{np} \end{pmatrix}$$

Thus if $p>1$, we immediately obtain a matrix A having a particular form.

Let F be a field, G a subfield of F and $F(x_1,\ldots,x_n)$ the field of rational functions in n variables with coefficients from F. Given $F\cup\Phi$, $\Phi = \{\phi_1,\ldots,\phi_t\}$, how many operations are required to compute $\Psi = \{\psi_1,\ldots,\psi_m\}$ using only the operations of addition, subtraction, multiplication and division?

Since F is given, and any operation within F results in an element in F, operations within F are not counted. Moreover multiplications or divisions, denoted by mult/div, by elements of the subfield G are not counted. The exclusion of these operations serves to strengthen the lower bounds. All results in this section are stated with the understanding that mult/div are counted in the above manner.

Call a set $\{v_1,\ldots,v_n\}$ linearly independent over G mod V if no nontrivial linear combination $c_1v_1+\ldots+c_nv_n$, with coefficients from G, is in V.

Denote by $S^{m\times n}$ the set of all $m\times n$ matrices with entries from a set S.

Theorem 1. Given $F\cup\Phi$, if $\Psi$ has r linearly independent elements over G mod $G\Phi+F$, where $G\Phi+F = \{c_1\phi_1+\ldots+c_t\phi_t+f \mid c_i \text{ in } G, f \text{ in } F\}$, then at least r mult/div are required to compute $\Psi$.

Proof. If an algorithm computes $\Psi$ with only s mult/div $\mu_1,\ldots,\mu_s$, where $\mu_j$ may depend on $\mu_i$ for $i<j$, then any element computed by the algorithm is of the form

$$b_1\mu_1+\ldots+b_s\mu_s + c_1\phi_1+\ldots+c_t\phi_t + f$$

with $b_1,\ldots,c_t$ in G and f in F. Define the column vectors $\psi'=(\psi_1,\ldots,\psi_r)$ $\mu=(\mu_1,\ldots,\mu_s)$ and $\phi=(\phi_1,\ldots,\phi_t)$, where $\psi_1,\ldots,\psi_r$ are the r linearly independent elements in $\Psi$. Thus there are matrices B in $G^{r\times s}$, C in $G^{r\times t}$ and d in $F^{r\times 1}$ such that

$$\psi' = B\mu + C\phi + d.$$

If $s<r$, the rows of B are linearly dependent, and a nontrivial vector a in $G^{1\times r}$ exists such that $aB = 0$. But then $a\Psi' = aC\phi+ad = c\phi+f$ for some vectof c in $G^{1\times t}$ and f in F, contradicting the hypothesis.

Corollary 1. Given $F\cup\{\xi^{i_1},\ldots,\xi^{i_t}\}$, if $\xi$ is not a root of any nontrivial polynomial of degree $\le n$ with coefficients from F, then at least m mult/div are required to compute $\Psi = \{\xi^{j_1},\ldots,\xi^{j_m}\}$ $1<j_1<\ldots<j_m\le n$ whenever $\{i_1,\ldots,i_t\}$ and $\{j_1,\ldots,j_m\}$ are disjoint.

For the linear case $\psi = A\phi$, where $\psi$ is the column m-vector $(\psi_1,\ldots,\psi_m)$, $\phi$ is the column t-vector $(\phi_1,\ldots,\phi_t)$ and A is a matrix in $F^{m\times t}$, Theorem 1 becomes:

Theorem 2. Given $F\cup\phi$, with the set $\phi = \{\phi_1,\ldots,\phi_t\}$ linearly independent over F mod F, if A has r linearly independent rows over G mod $G^{1\times t}$, then at least r mult/div are required to compute $A\phi$.

Proof. Let the first r rows of A be the linearly independent ones, A' be the submatrix of these r rows, and $\psi'$ be the r-vector $(\psi_1,\ldots,\psi_r)$.

As in Theorem 1, if $s<r$, there is a nontrivial vector a in $G^{1\times r}$ such that $a\psi' = aA'\phi = c\phi+f$ for some vector c in $G^{1\times t}$ and f in F, so that $(aA'-c)\phi=f$. But then, by the independence of $\{\phi_1,\ldots,\phi_t\}$, $aA'-c = 0$. This contradicts the independence of the rows of A'.

Corollary 2. Given $Fu\phi$, if A in $F^{m\times n}$ has r linearly independent rows over G mod $G^{1\times n}$, and the np entries of B in $\phi^{n\times p}$ are linearly independent over F mod F, then at least mp mult/div are required to compute AB.

Corollary 3. Given $Fu\{x_{11},\ldots,x_{np}\}$, with $F = G(y_{11},\ldots,y_{mn})$ at least mp mult/div are required to compute the product YX of the m×n matrix $Y = (y_{ij})$ and the n×p matrix $X = (x_{ij})$.

Theorem 3.(Winograd,[2]) Given the set $Fu\{x_1,\ldots,x_n\}$, if A in $F^{m\times n}$ has c linearly independent columns over G mod $G^{m\times 1}$, then at least c mult/div are required to compute Ax.

Corollary 4.(Winograd[2]) Given the set $G(y_1,\ldots,y_m) \cup \{x_{11},\ldots,x_{mn}\}$, at least mn mult/div are required to compute the product Xy of the m×n matrix $X = (x_{ij})$ and the vector $y = (y_i)$.

Theorem 4. Given $Fu\{x_1,\ldots,x_n\}$, if A has a submatrix S in $F^{r\times c}$, and there are no nontrivial vectors $\alpha$ in $G^{1\times r}$ and $\beta$ in $G^{c\times 1}$ such that $\alpha S\beta$ is in G, then at least r+c-1 mult/div are required to compute Ax.

Proof. Let A' be the r×n submatrix of A which contains S as a submatrix. If Ax can be computed with s mult/div, then there are

matrices B in $G^{r\times s}$, C in $G^{r\times n}$ and d in $F^{r\times 1}$ such that $A'x = B\mu + Cx + d$. Since $\alpha S\beta$ is not in G, the rows of A' must be linearly independent over G mod $G^{1\times n}$, so that $r\le s$ by Theorem 2. Partition B into $B = (M\ N)$, where N is an $r\times r-1$ matrix; then there is a nontrivial vector a in $G^{1\times r}$ such that $aN = 0$. Partition $\mu$ into $\mu = (\mu'\mu")$, where $\mu'$ is a column vector of $s-r+1$ elements; then

$$aA'x = aM\mu' + aCx + ad = a'\mu' + cx + f.$$

That is, there is a nontrivial linear combination aA', of the rows of A', such that aA'x can be computed with $s-r+1$ mult/div. But since $\alpha S\beta$ is not in G, aA' must have at least c linearly independent columns over G mod G. Thus aA'x requires at least c mult/div by Theorem 3, so that $s-r+1\ge c$.

Corollary 5.(Winograd,[3]) Given the set $R(y_1,y_2) \cup \{x_1,x_2\}$. At least 3 real mult/div are required to compute the product $(x_1 + ix_2)(y_1 + iy_2)$ of two complex numbers, (R= the reals).

Corollary 6. Given the set $R(y_1,\ldots,y_4) \cup \{x_1,\ldots,x_4\}$. At least 7 real mult/div are required to compute the product

$$(x_1 + ix_2 + jx_3 + kx_4)(y_1 + iy_2 + jy_3 + ky_4)$$

of two quaternions.

Theorem 5. Given the set $F\cup\{x_{11},\ldots,x_{mn}\} \cup \{y_1,\ldots,y_n\}$, if division is not allowed, then at least m(n-1) additions or subtractions are required to compute the product Xy of the $m\times n$ matrix $X = (x_{ij})$ and the n-vector $y = (y_i)$.

Proof. Let $\sigma(\Psi)$ be the minimum number of add/sub required to compute $\Psi$, and let $u = (1,1,\ldots,1)$. If division is not allowed, then $\sigma(Xu) \leq \sigma(Xy)$; whereas, if it is allowed, the algorithm for Xy may fail at y=u. The sum $s = x_{11}+\ldots+x_{mn}$ can be computed by first computing $s_i = (Xu)_i \quad i = 1,\ldots,m$ and then using m-1 more additions to compute $s = s_1+\ldots+s_m$. Hence $\sigma(s) \leq \sigma(Xu) + m-1$, and $\sigma(Xy) \geq \sigma(s)-m+1$. It can be shown [1] that $\sigma(x_1+\ldots+x_n) = n-1$. Thus $\sigma(s) = mn-1$, and $\sigma(Xy) \geq m(n-1)$.

## COMPLEXITY RELATIONS

When the underlying algebraic system is a ring, rather than a field, division may not be defined. In any event, unless a substantial reduction can be achieved by allowing division, algorithms in which division is not used are preferable. In the sequel, we assume that the only allowable operations are addition, multiplication and negation.

For any matrix A, denote by $\underline{A}$ the set of entries of A. Whenever we speak of computing Ab, we in fact mean computing the set $\underline{Ab}$. This fine point should be understood, since $\underline{A} = \underline{B}$ does not imply that A = B. Moreover, the matrix notation Ab simply serves to represent the set $\underline{Ab}$--nothing else is implied. Specifically, it is not necessary to have $\underline{A}$ and $\underline{b}$ in order to compute $\underline{Ab}$. For example, $a_0 + a_1x +\ldots+ a_nx^n$ can be represented as Ab, where $A = (a_0,\ldots,a_n)$ and b is the column vector $b = (1,x,\ldots,x^n)$. Horner's method may be used to compute $\underline{Ab}$, without computing $\underline{b}$.

Let $\pi(T|S)$ be the minimum number of multiplications required to

compute the set T give the set S, and $\sigma(T|S)$ be the minimum number of additions required to compute T given S. If $\chi$ stands for either $\pi$ or $\sigma$, the following relations hold.

Relation 1. $\chi(S|R) \leq \chi(T|R)$ if $S \subseteq T$.

Relation 2. $\chi(T|S) \leq \chi(T|R)$ if $R \subseteq S$.

Relation 3. $\chi(T|R) \leq \chi(T|R \cup S) + \chi(S|R)$.

Relation 1 is useful for obtaining lower bounds on $\chi(T|R)$, while 3 yields lower bounds on $\chi(S|R)$. Theorem 5 was proved by appealing to Relation 3 with $T = \{s\}$, $S = \underline{Xu}$ and $R = F \cup \{x_{11},\ldots,x_{mn}\}$.

In the sequel, we will confine our attention to computing Ab for worst case b; namely, when b is the column vector $x = (x_1,\ldots,x_n)$. Moreover, the entries of A are assumed to be independent of x.

Relation 4. Given S, if an independent variable in A or x is replaced everywhere by an element of S, to yield a new matrix B or a new vector y, then $\chi(\underline{By}|S) \leq \chi(\underline{Ax}|S)$.

This relation is useful for obtaining lower bounds on $\chi(\underline{Ax}|S)$. For example, if 1 is in S, then

$$\sigma(a_0+a_1x+\ldots+a_nx^n) \geq \sigma(a_0+a_1+\ldots+a_n).$$

When the given set S is fixed, since we are considering the worst case for x, we suppress x and S and simply write $\chi_A$ for $\chi(\underline{Ax}|S)$. Do not confuse $\chi_A$ with $\chi(\underline{Ax}|\underline{A},\underline{x})$.

Relation 5. $\chi_{AB} \leq \chi_A + \chi_B$.

Proof. $(AB)x = A(Bx)$.

Relation 5 gives rise to the following corollaries:

I. An elementary operation on A, can change $\chi A$ by at most $\pm 1$.

II. If $\chi P = 0$ and $\chi Q = 0$, then $\chi PAQ \leq \chi A$.

III. If A' is a submatrix of A, then $\chi A' \leq \chi A$.

IV. If $LP=I$, $QR=I$ and $\chi L=\chi P=\chi Q=\chi R=0$, then $\chi PAQ = \chi A$.

V. Write $A' \sim A$ if A' can be obtained by permuting the rows or columns of A. If $A' \sim A$, then $\chi A' = \chi A$.

VI. If A' is obtained from A by addition/subtraction of one row (column) to another row (column), then $\pi A' = \pi A$.

VII. For $n \geq 0$, $A^n \leq n\chi A$.

VIII. If $A = \begin{pmatrix} B \\ CB \end{pmatrix}$, then $\chi A \leq \chi B + \chi C$.

IX. If $A = (B\ BC)$, then $\pi A \leq \pi B + \pi C$.

X. Given F, if A in $F^{m\times n}$ has rank r, then $\pi A \leq r(m+n-r)$ and $\sigma A \leq r(m+n-r)-n$.

Relation 6. $\chi A\pm B \leq \chi A+\chi B+\chi(\underline{Ax\pm Bx}|\underline{Ax},\underline{Bx})$.

Proof. $(A+B)x=Ax\pm Bx$ and $(-B)x=-(Bx)$.

Relation 6 gives rise to the following corollaries:

XI. $\pi A\pm B \leq \pi A+\pi B$.

XII. $\chi A \oplus B \leq \chi A + \chi B$. $\quad A \oplus B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$

XIII. The Kronecker product $A \times B = (a_{ij}B)$ is obtained by replacing $a_{ij}$ in A by $a_{ij}B$. If A is $m \times n$ and B is $r \times s$, then $A \times B$ is $mr \times ns$, and $\chi A \times B \leq r\chi A + n\chi B$.

XIV. If $A^{[r]}$ is the $r^{th}$ Kronecker power of A, and A is $m \times m$, then $A^{[r]}$ is $n \times n$, $n = m^r$, and $\chi A^{[r]} \leq \frac{n}{m}(\log_m n)\chi A$

Notice that by Theorem 2, given the set $G(y_{11},\ldots,y_{mn}) \cup \{x_{11},\ldots,x_{rs}\}$, at least mnrs mult/div are required to compute $\underline{Y \times X}$.

We close this section with the two following observations:

If A is in $Z^{m \times n}$ (the integers), then $\pi A = 0$, since every multiplication by an integer constant can be replaced by additions.

If A is in $Q^{m \times n}$ (the rationals), then $\pi A \leq \min(m,n)$, since $A = d^{-1}B = Bd^{-1}$, where d is a common denominator of the entries of A, and $B = dA$ is in $Z^{m \times n}$.

We note in passing, that $\pi A$ and $\sigma A$ are two independent measures, and that there may exist no single algorithm which simultaneously achieves both $\pi A$ and $\sigma A$. For example, if $A = [k]$ is a $1 \times 1$ matrix with $k > 0$ an integer constant, then $\sigma A = 0$ since $Ax = kx$. On the other hand, $\pi A = 0$ since $Ax = x + \ldots + x$.

## UPPER BOUNDS

In view of Corollary 4 and Theorem 5, the usual algorithm for matrix multiplication achieves both lower bounds. However, these bounds apply

only to the worst case Xy. In many cases, we are interested in computing Ax, where A has some particular form. In this section we examine several such matrices. Unless otherwise specified, all multiplications are counted.

SYMMETRIC-LIKE MATRICES

Using matrices, the identities

$$\varepsilon au + \varepsilon av = \varepsilon a(u+v) \qquad au + \varepsilon av = a(u+\varepsilon v)$$

$$au + av = a(u+v) \qquad \varepsilon au + av = \varepsilon a(u+\varepsilon v)$$

where $\varepsilon = \pm 1$, may be expressed as

$$B=\begin{pmatrix} \varepsilon a & \varepsilon a \\ a & a \end{pmatrix}=\begin{pmatrix} \varepsilon \\ 1 \end{pmatrix} a \begin{pmatrix} 1 & 1 \end{pmatrix} \qquad C=\begin{pmatrix} a & \varepsilon a \\ \varepsilon a & a \end{pmatrix}=\begin{pmatrix} 1 \\ \varepsilon \end{pmatrix} a \begin{pmatrix} 1 & \varepsilon \end{pmatrix}.$$

Corollaries II and III show that $\pi B=\pi C=\pi a$. Then

$$A=\begin{pmatrix} a & \varepsilon b \\ b & c \end{pmatrix}=\begin{pmatrix} a-\varepsilon b & 0 \\ 0 & c-b \end{pmatrix}+\begin{pmatrix} \varepsilon b & \varepsilon b \\ b & b \end{pmatrix}$$

$$\pi A \le \pi(a-\varepsilon b) + \pi(c-b) + \pi b$$

The 3×3 case will suggest a general technique:

$$A=\begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}=\begin{pmatrix} u & 0 & 0 \\ 0 & v & 0 \\ 0 & 0 & w \end{pmatrix}+\begin{pmatrix} b & b & 0 \\ b & b & 0 \\ 0 & 0 & 0 \end{pmatrix}+\begin{pmatrix} c & 0 & c \\ 0 & 0 & 0 \\ c & 0 & c \end{pmatrix}+\begin{pmatrix} 0 & 0 & 0 \\ 0 & e & e \\ 0 & e & e \end{pmatrix}$$

where $u = a-b-c$, $v = d-b-e$ and $w = f-c-e$.

$$\pi A \le \pi u + \pi v + \pi w + \pi b + \pi c + \pi e$$

Proposition 1. Given A, if A is n×n and of the form $a_{ij} = \pm a_{ji}$, then $\pi A \le n(n+1)/2$, even if multiplication is not commutative.

Corollary 7. Two complex numbers can be multiplied with 3 real multiplications.

Corollary 8. Two quaternions can be multiplied with 10 real multiplications.

This technique is applicable to any matrix $B \sim A$. Moreover, the entries of A may themselves be matrices.

An important symmetric-like matrix, with numerous applications, is the Toeplitz matrix, defined by $a_{ij} = a_{i+1,j+1}$. The 4×4 Toeplitz matrix is given by

$$T_4 = \left(\begin{array}{cc|cc} d & c & b & a \\ e & d & c & b \\ \hline f & e & d & c \\ g & f & e & d \end{array}\right) = \begin{pmatrix} B & C \\ D & B \end{pmatrix}$$

The important thing to observe is that if $n = rs$, then $T_n$ can be partitioned into an r×r Toeplitz matrix whose entries are s×s Toeplitz matrices. Moreover, if A and B are m×n Toeplitz matrices, then A±B can be computed with m+n-1 add/sub, and A±B is itself Toeplitz.

Corollary 9. Given $T_n$, if $n = rs$, then $\pi T_n \leq r(r+1)/2\ \pi T_s$. Using this recursion, if $n = n_1 \ldots n_k$, then $\pi T_n \leq \prod_1^k n_i(n_i+1)/2$. When $n = 2^k$, then $\pi T_n \leq 3^k = n^{\log_2 3} \approx n^{1.58}$. Since $T_n$ can always be viewed as a submatrix of $T_{2^k}$ whenever $n \leq 2^k$, if $\lceil x \rceil$ denotes the smallest integer $\geq x$, we have:

Proposition 2. Given A, if A is an n×n Toeplitz matrix, then $A \leq 3^{\lceil \log_2 n \rceil}$, even if multiplication is not commutative.

Corollary 10. Two sequences of lengths m and n can be convolved with $3^{\lceil \log_2 t \rceil}$ multiplications, t = m+n-1.

Corollary 11. Two polynomials of degrees m and n, can be multiplied with $3^{\lceil \log_2 t \rceil}$ multiplications, t = m+n+1.

Corollary 12. Two numbers of m and n digits can be multiplied with $3^{\lceil \log_2 t \rceil}$ multiplications, t = m+n-1.

Corollary 13. Given A, if A is an n×n circulant, defined by $a_{ij} = a_{i+1,j+1 \pmod n}$, then $\pi A \leq 3^{\lceil \log_2 n \rceil}$.

In all cases, multiplication need not be commutative. Moreover, given S, all results are contained in the smallest ring which contains S.

COMPANION MATRICES

In this subsection, we assume that the underlying algebraic system is a field F.

The companion matrix of the polynomial

$$\phi(x) = x^n - c_{n-1}x^{n-1} - \ldots - c_0$$

is the n×n matrix

$$C_\phi = \begin{pmatrix} 0 & & & c_0 \\ 1 & \ddots & & c_1 \\ & \ddots & & \vdots \\ & & 1 & c_{n-1} \end{pmatrix}$$

If D is the n×n diagonal matrix with entries $\alpha_0,\ldots,\alpha_{n-1}$, and u is the column n-vector u = (1,...,1), the Vandermonde matrix $V = (\alpha_r^s)$, generated by $\alpha_0,\ldots,\alpha_{n-1}$, is the n×n matrix

$$V = (u\ Du\ \ldots\ D^{n-1}u).$$

Shift Theorem. If $\phi(\alpha_r) = 0$ for r=0,...,n-1, then $DV = VC_\phi$. If $\alpha_0,\ldots,\alpha_{n-1}$ are distinct, then $V^{-1}$ exists, $D = VC_\phi V^{-1}$, and $p(D) = Vp(C_\phi)V^{-1}$ for any polynomial p(x).

Let P be the column vector $(a_0,\ldots,a_{n-1})$ associated with the polynomial

$$p(x) = a_0 + a_1x + \ldots + a_{n-1}x^{n-1}.$$

It is easily verified that if P is associated with p(x), then $C_\phi P$ is associated with $xp(x) \bmod \phi(x)$. Moreover, if

$$q(x) = b_0 + b_1x + \ldots + b_{n-1}x^{n-1}$$

then $q(C_\phi)P$ is associated with $q(x)p(x) \bmod \phi(x)$.

If $\alpha_0,\ldots,\alpha_{n-1}$ are distinct, then $V^{-1}$ exists and

$$q(C_\phi)P = V^{-1}Vq(C_\phi)V^{-1}VP = V^{-1}q(D)VP.$$

If Q is associated with q(x), the elements of VQ are identical to the diagonal entries of q(D). Thus q(D)VP is the term-by-term product (VQ)·(VP) of the two column vectors VQ and VP. Then

$$q(C_\phi)P = V^{-1}[(VQ)\cdot(VP)]$$

Call VQ the transform of Q. Then $q(C_\phi)P$ can be computed as follows:

1. compute VP.
2. compute VQ.
3. multiply VQ by VP term-by-term.
4. take the inverse transform.

If the degree of $q(x)p(x)$ is less than n, then

$$q(x)p(x) = q(x)p(x) \bmod \phi(x).$$

Thus the process of computing the coefficient vector $q(C_\phi)P$ of $q(x)p(x) \bmod \phi(x)$ may be referred to as convolution mod $\phi(x)$. Thus we have:

Proposition 3. Any nonsingular Vandermonde matrix V has the convolution property.

Choose distinct integers $\alpha_0,\ldots,\alpha_{n-1}$, then $\pi V = 0$. Moreover, the entries of $V^{-1}$ are rational constants, thus $V^{-1} = k^{-1}U$ for some integer k and integer matrix U. It follows that $U(VQ)\cdot(VP)$ can be computed with n multiplications. Multiplication by $k^{-1}$ takes no more than n multiplications.

The process described above can be performed in any field of characteristic zero, or in any field of sufficiently large characteristic. Thus we have:

Proposition 4. In a field of characteristic zero or $p \geq n$, if $\phi(x)$ has distinct integer roots, then $q(x)p(x) \bmod \phi(x)$ can be computed with 2n multiplications.

Corollary 14. In a field of characteristic zero or $p \geq m+n-1$, two sequences of lengths m and n can be convolved with $m+n-1 + \min(m,n)$ multiplications.

Corollary 15. In a field of characteristic zero or $p \geq m+n+1$ two polynomials of degree m and n can be multiplied with $m+n+1 + \min(m+1,n+1)$ multiplications.

Corollary 16. In a field of characteristic zero or $p \geq m+n-1$, given A, if A is an $n \times n$ circulant, then $\pi A \leq 3n-1$.

For prime n, computing the finite Fourier transform can be shown to be equivalent to multiplying by an $(n-1)\times(n-1)$ circulant [5].

Corollary 17. For prime n, the finite Fourier transform can be computed with $3n-4$ multiplications.

We close with a conjecture. If C is the companion matrix of $x^n-1$, any $n \times n$ matrix A can be written as

$$A = A_0 + A_1 C + \ldots + A_{n-1} C^{n-1}$$

where $A_0,\ldots,A_{n-1}$ are $n \times n$ diagonal matrices. Then

$$AB = (A_0 + \ldots + A_{n-1} C^{n-1})(B_0 + \ldots + B_{n-1} C^{n-1})$$

appears as a cyclic convolution, except that C does not commute with the coefficients. Nevertheless, based on the above results, we make the following

Conjecture. Two $n \times n$ matrices can be multiplied with about $2n^2$ multiplications!

## REFERENCES

[1] Ostrowski, A.M., "On two problems in abstract algebra connected with Horner's rule," Studies presented to R. vonMises, Academic Press, New York, 1954, pp. 40-48.

[2] Winograd, S., "On the number of multiplications necessary to compute certain functions," Comm. Pure and Applied Math., vol. 23, 1970, pp. 165-179.

[3] Winograd, S., "On the multiplication of 2×2 matrices," IBM Research Report RC267, Jan. 1970.

[4] Cooley, J.W. and J.W. Tuckey, "An algorithm for the machine computation of complex Fourier series," Math. Computation, vol. 18, April 1965, pp. 297-301.

[5] Rader, C.M., "Discrete Fourier transforms when the number of data samples is prime," Proc. IEEE, vol. 56, June 1968, pp. 1107-1108.

[6] Fiduccia, C.M., "Fast matrix multiplication," Doctoral dissertation (being completed), Brown University.